Using AICc

The Akaike Information Criteria (AIC) is a key part of "the new statistics." The fundamental goal: find the model – among your list of hypothesized alternatives – that is most plausible. AICs can be applied to categorical predictors (as used in ANOVAs), continuous predictors (as used in regression), or combinations of both.

AIC discounts models for the number of variables to find the **most plausible model**. Multiple R packages report AIC metrics, including bbmle and AICcmodavg, which produce simple tables to compare models. Here we use bbmle because it is simple to code. Reported metrics include:

- AIC or corrected AIC (AICc). The AICc should be your default, because it corrects for low N and equals AIC at large N. Lower values indicate more plausible models.
- delta AICc. The difference between ranked models. A delta AICc ~ 2 indicates a clear choice – otherwise, two models are comparable.
- AICc weight (w_i). This represents the relative likelihood of a model, where 1.0 = most likely. ***Weight is the best way to rank and compare models***.

Load and attach our copter data from:
https://sciences.ucf.edu/biology/d4lab/wp-content/uploads/sites/23/2021/09/helicopter-data.csv

And make fold, wing, and group factors for categorical treatments (ANOVA-style output):

```
data$ffold<- factor(Fold)
data$fwl <- factor(WL)
data$fbw <- factor(BW)
data$fgroup <- factor(Group)
```

The first, important step in using AICs is to carefully construct hypotheses to test. **AIC only compares models you list** – you have to think first about which ones you are interested in, for good reasons. Consider these:

```
nullmodel <- lm(Time ~ 1) # This says Time is a constant; nothing in the
   # experiment mattered.
   # ALWAYS INCLUDE NULLS IN YOUR COMPARED MODEL SETS!!!
gsmodel <- lm(Time ~ fgroup + Step) # “only groups & steps mattered”
wgmodel <- lm(Time ~ fwl + fgroup) # “only wings mattered”
fgmodel <- lm(Time ~ ffold + fgroup) # “only folds mattered”
wfgsmodel <- lm(Time ~ fwl*ffold + fgroup + Step)
fullmodel  <- lm(Time ~ Step + fwl + fbw + ffold + Group + fwl:fbw +
    fwl:ffold + fbw:ffold) # Our full, hypothesized experimental design, as
    # simplified in our previous class
```

Not all possible model combinations are listed? Why? Because these are all ***I*** hypothesized. You want more? Go ahead! But choose wisely: To throw all possible models at a question is like fishing with dynamite – you may get something, but there are consequences – a main one being you didn’t do Science (i.e., test ideas).

To compute AICs, install (if not already done) and load the **bbmle** package.
Alternatively, install and load the **AICcmodavg** package.
As with much of R, there are many ways to get the same result.

For **bbmle**: run this command to generate a table for AICc scores, etc.:

```
AICctab(nullmodel, gsmodel, wgmodel, fgmodel, wfgsmodel, fullmodel, base=T,
delta=T, weights=T)
# This simply asks for table of listed models, including base AICc values,
delta AICc values, and weights
```

Or for **AICcmodavg**, run these commands:

```
hypothmodels <- list (nullmodel, gsmodel, wgmodel, fgmodel, wfgsmodel,
fullmodel)
Cand.names <- c("nullmodel", "gsmodel", "wgmodel", "fgmodel", "wfgsmodel",
"fullmodel")
aictab(hypothmodels)
```

To be clear, we already expect our simplified full model to be most plausible – we used stepIAC last class to find that one.

***See the weights*?** Are justified in presenting the full ANOVA and lm results?

```
summary(fullmodel)
summary.aov(fullmodel)
```

<u>Now let's use AICctab for another data set.</u>

Get the data set:

1. In the MASS package, there is a data set on 1993 cars. If MASS is not already installed, install it now. If MASS is already installed, then turn it on by clicking the box or:

```
library(MASS)
```

2. Because it comes with a package, we load Cars93 differently than if when we import a txt file:

```
data(Cars93)
attach(Cars93) # notice that this replaces copters data
```

I start with two hypotheses and compare them with AICc – as a template to show you how to proceed. **Then you make three more models and compare them all.**

<u>Bet 1</u>: I bet MPG.city is simply predicted by Origin (US vs non-US cars), because in 1993 ‘Murican cars were bigger. Thus my model 1 and request for output looks like this:

```
model1 <- lm(MPG.city ~ Origin)
summary(model1)
summary.aov(model1)
plot(MPG.city ~ Origin)
```

Bet 2: I bet MPG.city in 1993 can be simply predicted by Manufacturer, because the Big Three (Ford, Chevy, Chrysler and their subsidiaries - e.g., Cadillac, Lincoln) made big cars.
Adapt model1 code above for Bet 1 to run a model2 for Bet 2.

Examine the Adjusted R^2 of the two models using `summary(model1)` and `summary(model2)`
Which model would you think is the best?
Now compare those two models using AICc instead:

```
AICctab(model1, model2, weights = TRUE, delta = TRUE, base=TRUE, sort = TRUE)
```

So... the model based on Manufacturer has a lower p value, but the Origin-based model more parsimoniously explains MPG.city. AIC-based model selection is not p-value based. Instead, the most *efficient* models are most plausible. It assesses "bang for the buck" – more complex models are often more plausible (and should be if terms are useful), but added terms are discounted and so AIC evaluates a complexity/value-added tradeoff for you.

What other factors might also affect MPG.city? **Construct at least THREE MORE alternative models** to evaluate by adding terms to model1 above. Make models as complex as you think is required, BUT a model should represent a hypothesis – such as my bets above. *Grab-bag / smörgåsbord / all-possible-options models do not count because this approach is about testing* **a priori** *hypotheses*. Evaluate each model (as above), and then compare all the models using `AICctab`. Include a null model in your comparisons!

Now look at the Adjusted R^2 of the models you evaluate – this should be used to "criticize" the "best" model - because the best model may still be lousy. R^2 is helpful for this purpose, *but it is not a fair way to compare models with different explanatory variables*. And notice that you should report model coefficients, adjusted R^2 values, and other output for your most-plausible model(s) identified by AICc.

Guidelines for using AICs, etc.:
- Models represent your operational hypotheses – think and specify clearly.
- Use AICs to select among models – not R^2 values.
- Models less plausible than a null can be set aside
- If a model has the greatest AIC weight and the next-ranked model has δAIC > 2, then one model is clearly most plausible.
- If δAIC $\leq$ 2, then one model is not clearly most plausible. Examine details; if models are nested (e.g., Y ~ X1 + X2 vs. Y ~ X1), then consider outputs to decide if X2 should be reported. Otherwise, use model averaging
- After selecting your most-plausible model(s), describe the model(s) with output details,

including R^2.

- Also, report its compliance with model assumptions (use the performance package to make that easy!).